

A complete guide to

/plog/
C H I P S P E E C H

Manual v1.7

Welcome to Plogue chipspeech

It said in a menacing tone: "Another Visitor... Stay a while. STAY FOREVER!". I stood there in awe, turning around at my school buddy:

- 'Did it just talk?'
- 'Wait until you hear the rest', my friend replied.

That was it.

Speech synthesis had just left a permanent mark in my mind, and has fascinated me ever since. Though I've been doing research in this field for quite a while now, I still get chills each time I unearth a hidden word from a vintage device. To me, it just feels like a buried time capsule is suddenly revealed.

Speech synthesis came a long way since those early days, but in some ways, it lost some of its charms. Vintage speech chips were not *quite* understandable. Each had its own unique textures, twists, and artifacts. Yet somehow, their limited capabilities can sound 'musical' to anyone with a "synthesizer ear".

Our goal with chipspeech was to preserve the unique textures of these voices, while at the same time giving them advanced musical flexibility. You can now type in any English phrase and play anything on your MIDI Keyboard or through your DAW, and chipspeech will do the rest.

But since none of this would have been possible without the ingenuity of all the pioneers in the field of speech synthesis, we have decided to dedicate this product to them:

Homer Dudley, Forrest S. Mozer, Denis Klatt, John Larry Kelly, Jr , Gene A. Frantz, Larry Brantingham and the others at Texas Instruments, and all the others we may have forgotten.

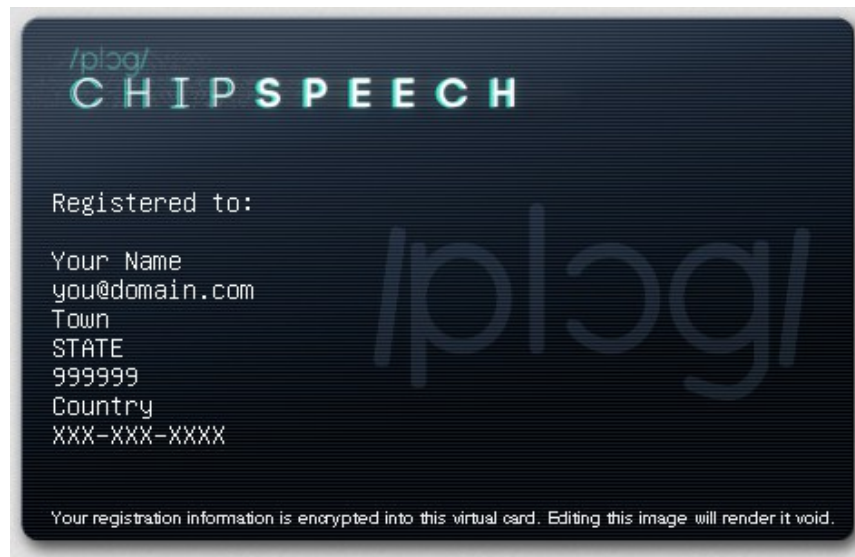
Thank you!

David Viens
January 2015

Authorizing chipspeech

Plogue chipspeech needs to be authorized in order to be fully functional, otherwise it will run in **DEMO** mode. When you order a license from Share*it, you receive a personal Activation Key card named **aria_key_1017.png** as an email.

The Activation Key card is an image resembling a typical credit card. This image contains your registration and details encoded within the Key card image. It will look like this:



You should save the **aria_key_1017.png** image file to your hard drive (keep it in a safe place). For convenience, we recommend that you initially save the .png file to your desktop. You will also receive a copy of the license card in your email inbox.

- 1) Locate the “license card” image where you saved it on your hard drive. (**aria_key_1017.png**)
- 2) Open the chipspeech software application, or launch your favorite host and make sure you see chipspeech interface.
- 3) Simply click and hold on the file, drag the “license image” or file icon directly onto the application's UI itself, and release it.

If you don't get any message (or are not able to drop the key in that host), try to **import** the png file from the the snapshot load menu or the slot load import. (use *.* as file filter)

You should be presented with a message saying “Plogue Art et Technologie, Inc chipspeech is now activated for {your name here}”

If none of those methods work, please contact chipspeech.support@plogue.com and attach your key.

Extremely Important!!

The **aria_key_1017.png** file contains your **sensitive personal information**, encrypted inside, including your **full name and address** taken from the online shop. Carefully protect this file. **DO NOT GIVE THIS FILE TO ANY-ONE OR DISTRIBUTE IT IN ANY WAY OR YOUR PERSONAL INFORMATION WILL BE COMPROMISED. IF THE FILE BECOMES PUBLIC THE CARD NUMBER WILL BE BLACKLISTED AND THE CARD REVOKED. WE ARE NOT RESPONSIBLE IF YOU GIVE YOUR PERSONAL DETAILS TO A THIRD PARTY. IF THE CARD IS STOLEN, CONTACT US IMMEDIATELY.** Without a valid card you will also not be able to obtain critical updates to the program.

Important Note:

If you have special circumstances or require site licensing, please contact us.

Updating to the Latest Version

Be sure to check the Plogue Web sites for any possible updates that have occurred since the time your version of the software was released. Software is frequently updated and a more recent version may be available.

Getting Help

The first place to look for a solution to any problem you may be experiencing is in this manual. Please read the manual before contacting support. Next, check the readme files (if any) which contain important information and all last-minute changes that haven't been available when creating this guide.

The Plogue chipspeech Player and sounds are dynamic — evolving and growing. Please check the support area of our website at www.Plogue.com for the latest up-to-date information products, troubleshooting, FAQs, helpful hints and tutorials. Another resource is the support forums.

Whenever you encounter problems, you should also check if you have installed the latest updates. The version number of your software is displayed in the **About** dialog. Updates are released regularly to fix known problems and to improve the software.

If you can't find a solution to your problem please email us at chipspeech.support@plogue.com. The best way to get the help you need is by giving us plenty of detailed information about the problem you are having. We do ask you to read this guide thoroughly and exhaust the other avenues of support before contacting us.

The Plogue forum can be accessed at: <http://www.plogue.com/phpBB3/>

You don't have to register to browse posts, but before you can post, you will have to sign up.

End User License Agreement

See Licence.rtf, which contains the full license agreement.

About this Manual

The information in this document is subject to change without notice and does not represent a commitment on the part of Plogue Art et Technologie, Inc. No part of this publication may be copied, reproduced or otherwise transmitted or recorded, for any purpose, without prior written permission by Plogue Art et Technologie, Inc.

Plogue chipspeech® is a registered trademark of Plogue Art et Technologie, Inc. Use of Plogue chipspeech® software instrument and the contents herein are subject to the terms and conditions of the license agreement distributed with the program (see license.rtf). You should carefully read the license agreement before using this product. The sounds presented in Plogue chipspeech® are protected by copyright and cannot be distributed, whether modified or unmodified. The Guide to Plogue chipspeech contained herein are also covered by copyright. ARIA™ is a trademark of MakeMusic Inc and Plogue Art et Technologie Inc. The ARIA Engine is copyrighted by Plogue Art et Technologie and MakeMusic Inc. Any other trademarks of third-party programs are trademarks of their respective owners. The information contained herein may change without notice and does not represent a commitment on the part of Plogue Art et Technologie, Inc.

Plogue Art et Technologie, Inc.
www.plogue.com
www.facebook.com/wwwPLOGUEcom
twitter.com/wwwPLOGUEcom

Chipspeech Quick Usage Guide

Using chipspeech

- Enable full keyboard input. You will need this to type in your lyrics!
- Chose a voice and edit its parameters to your liking in the *Controls* tab. You will also find more advanced parameters in the *Modulation* tab. See the 'Parameter Reference' section.
- Type the lyrics to your song in the *Words* tab.
- Play in the melody.
- Enjoy!

Activating full keyboard input

Many plug-in hosts will intercept key strokes before they get to plug-ins, and have an “enable keyboard input” option to disable this feature. Since chipspeech is based around text input, you will probably need activate this option.

Host	Location
Modplug	Plugin window menu: Options->Pass Keys to Plugin
Reaper	Right-click the plug-in instance, check “Send all keyboard input to plug-in” option
Renoise	Bottom of plug-in window: “Enable keyboard” button
Sonar	Top of plug-in window: Keyboard shaped button

Controlling the playback cursor

Plogue chipspeech keeps a cursor in the text, and advances by one syllable for each MIDI note. Because the song can be played starting from any position, the text cursor may end up getting out of sync with the song.

- Midi CC2 will set the cursor to a text line. For instance, if chipspeech receives CC2 value 57, the text cursor will be set at the beginning of line 57 for the next note.
- Midi patch change will do the same thing as CC2 (provided for convenience for some Midi controllers).
- Key switches (low notes from 0 to 23) will trigger lines 0 to 23. (higher lines cannot be triggered with key switches)
- Editing or selecting a text line will do the same, and set the next playback start position to that line as well.
- Every time playback is reset, the text cursor will restart at the last edited text line.
- Midi CC3 will set the playback cursor within the current line. (not available in Talking mode)
- In Talking mode, the cursor advances line-by-line instead of syllable-by-syllable.

English text input

Plogue chipspeech contains an English phonetic dictionary (CMUDict) and will automatically generate a pronunciation for any English text that you type in. Furthermore, it will try to guess the pronunciation of unknown words. All you have to do is to type in the lyrics in the *Words* tab.

- The copy/paste/undo buttons apply to the *whole song*. They let you import and export your lyrics to other programs. The “undo” button only works for a single undo step.
- Copy/paste/undo keyboard shortcuts (Ctrl+C/V/Z) only apply to the *currently edited line*.
- If you don't like the pronunciation of a word (for instance, you need the 3 syllable ev-e-ry instead of the default 2 syllable eve-ry), the pronunciation can be edited in X-SAMPA section. X-SAMPA words are shown between brackets in the English section - for instance, 3 syllable ev-e-ry will be shown as [E v@ ri.
- Periods (“.”) will be translated into pauses ([#] in X-SAMPA).
- Words can be extended into multiple syllables by the use of dashes (“-”). Example: the word “che-e-eck” will be translated as [tSE E Ek] in X-SAMPA (plays over 3 syllable instead of one). (note: not supported by talking mode)
- Chipspeech also features limited Japanese support: hiragana and katakana text will be translated into the corresponding pronunciations (kanji is not supported though). (note: not supported by talking mode)

Phonetic text input

Chipspeech uses X-SAMPA for phonetic transcription. This was chosen to eventually allow for multiple language support (unlike Arpabet or Votox), and to remove the need for non-ASCII character support (which is a problem with IPA and Americanist notation). See the chipspeech X-Sampa reference for more info. Note: Talker voices do not support phonetic text input.

Using Circuit Bending Voices

Circuit Bending voices let you play vintage speech data through a bit accurate emulation, just as it would have been on the original hardware, while messing up the playback creatively by changing the pitch, data offsets and loop points.

On the TMS chip, the jump at the end of the loop will land into random locations and will cause strange, half-scrambled recombinations of the synthesis data, causing the oscillator/noise and filter to go through all sorts of weird settings, rings and pops and whatnot. This is what makes Speak&Spell circuit bending so rich and interesting : by shorting the address pins of the speech data ROM, you would can create strange, half-coherent loops. Plogue chipspeech lets you recreate this sound very easily : simply pick the “TMS Bending” voice, select a bank and an entry, and then mess up with the loop start and loop length settings to your heart's content!

Using Talker Voices

Talker voices let you use emulations of whole speech synthesis systems, all the way from text-to-phoneme conversion, speech melody and rhythm generation, to sound production, which makes it easy to get the full original tonality of some iconic voices like CiderTalk'84 and Rotten.ST. These voices do not support X-SAMPA or Japanese input, and in many cases will parse the input differently from Chipspeech. They are also sensitive to punctuation, and will give different inflections. Not all voice synthesis parameters will work on these voices.

Meet the Singers!



Bert Gotrax

Inventor: Richard T. Gagnon

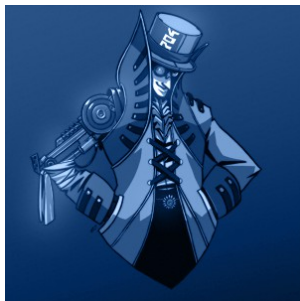
Famous for: Q-Bert, Gorf

Device: Votrax SC-01

Wikipedia: <http://en.wikipedia.org/wiki/Votrax>

Synthesis type: Formant (analog)

Tone: Lots of complex frequency whines and noises, due to real analog hardware. May benefit from some EQ (especially increased bass and treble and lowered mids).



Dandy 704

Inventors: John Kelly and Carol Lochbaum

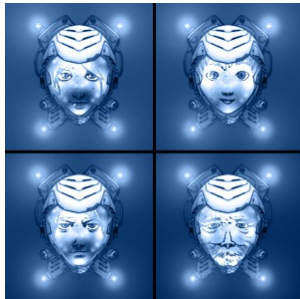
Famous for: Daisy Bell (first computer voice synthesis song)

Device: System based on an IBM 704 computer

Wikipedia: http://en.wikipedia.org/wiki/Daisy_Bell#In_technology_and_culture

Synthesis type: Physical modelling (waveguide)

Tone: Dark and muted, many strange inflections in the vowels due to the tortuous signal path.



Dee Klatt

Inventor: Dennis Klatt

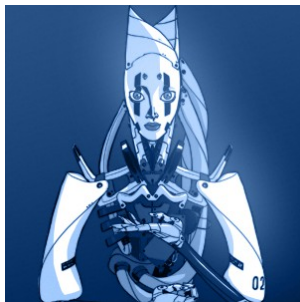
Famous for: Stephen Hawking's voice

Device: Dectalk

Wikipedia: <http://en.wikipedia.org/wiki/DECtalk>

Synthesis type: Formant

Tone: Very versatile. This voice can morph all the way from husky male baritones up to high female divas.



Lady Parsec

Inventor: Richard Wiggins & the Speak & Spell team

Famous for: Parsec (TI-99/4A game)

Device: TI-99/4A plug-in speech synthesizer module

Wikipedia: [http://en.wikipedia.org/wiki/Parsec_\(video_game\)](http://en.wikipedia.org/wiki/Parsec_(video_game))

Synthesis type: LPC10 (lattice)

Tone: Muted and warbly.



Lady Parsec HD

Inventor: The Plogue team
Device: Your computer
Synthesis type: Plogue's "Formant Singer" technology.
Tone: Clear and "human".



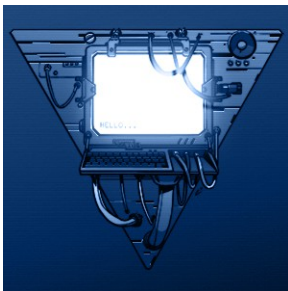
Otto Mozer

Inventor: Forrest Mozer
Famous for: Berzerk
Device: TSI
Wikipedia: http://en.wikipedia.org/wiki/ESS_Technology
Synthesis type: Voice segment sequencing
Tone: Very bright and crunchy, this assertive voice will pierce through any mix. Its bass range is extremely fat.



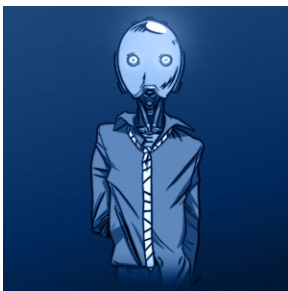
Spencer AL2

Patent holder: John E. Stork
Famous for: Intellivoice (Intellivision), Magnavox Odyssey² voice attachment
Device: SP0256-AL2 chip
Wikipedia: http://en.wikipedia.org/wiki/General_Instrument_SP0256
Synthesis type: LPC12 (cascaded 2-poles)
Tone: Very choppy due to the limited number of voice segments and lack of interpolation.



Terminal 99

Inventor: Richard Wiggins & the Speak & Spell team
Famous for: TI-99/4A Terminal Emulator II cartridge
Device: TI-99/4A plug-in speech synthesizer module
Wikipedia: http://en.wikipedia.org/wiki/Texas_Instruments_TI-99/4A
Synthesis type: LPC10 (lattice)
Tone: Somewhat choppy.



VOSIM

Inventors: Werner Kaegi and Stan Tempelaars
Device: Standard DAC
Wikipedia: [http://en.wikipedia.org/wiki/Werner_Kaegi_\(composer\)](http://en.wikipedia.org/wiki/Werner_Kaegi_(composer))
Synthesis type: Windowed Oscillator Sync
Tone: Synthesizer-like.



CiderTalk'84

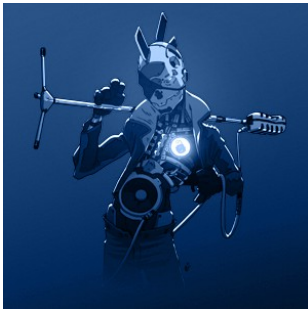
Inventors: [Softvoice, inc.](#)

Famous for: Apple MacInTalk and Amiga narrator.device

Wikipedia: http://en.wikipedia.org/wiki/Speech_synthesis#Apple

Synthesis type: Windowed Oscillator Sync

Tone: Crunchy and somewhat lively.



Rotten.ST

Inventors: Martin Day and Andy Beveridge

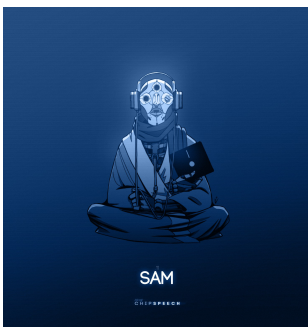
Famous for: U96 - Das Boot, Quadrophonia, other early Techno

Device: Atari ST

Wikipedia: http://en.wikipedia.org/wiki/Speech_synthesis#Atari

Synthesis type: Windowed Oscillator Sync

Tone: Buzzy and aggressive.



SAM

Inventor: [Softvoice, inc.](#)

Famous for: Can you repeat the question?

Device: Any 4-bit DAC

Wikipedia: https://en.wikipedia.org/wiki/Speech_synthesis#Apple

Synthesis type: Formant synthesis.

Tone: Deep, rich and metallic with an unplaceable accent, spanning a wide range of frequencies.



Voder

Inventor: Homer Dudley

Famous for: Being the first electro-mecanic voice synthesizer

Device: Rigorously trained human operator

Wikipedia: <http://en.wikipedia.org/wiki/Voder>

Synthesis type: Filter bank synthesis

Tone: Strange resonances with unexpected surprises and weird noise, can be very muted.

Parameter Reference

Global Parameters



Voice : Lets you select which voice to use. Voices are divided into 2 categories:

- Voice : Singing voice using lyrics. This is intended for musical use.
- Bending : Circuit bending. This lets you playback and loop raw preset voice data, with often interesting but unpredictable results.

Snapshot : Snapshots are “Master Programs”, complete setups with a “snapshot” of the current state of the system at the moment of saving, including parameter values, MIDI controls and lyrics. You can save any number of snapshots into your user folder and even create sub-folders, which will mimic the structure of the the default programs. Snapshots have the .ariax extension.

Pitch Bend Range : Maximum amount of effect the MIDI pitch bender has on the pitch, in semitones.

Helium / Sample Rate Bend Range : Maximum amount of effect the MIDI pitch bender has on the voice resonant frequencies (formants), in semitones.

Mono/Poly : Switches between single voice (lets you play legatos) and multiple voices (lets you play chords).

Volume : Overall volume level of the voice.

Phoneme Speed : Changes how fast each vowel transition is played, without changing the overall playback rate.

Wave Rate : How fast the voice plays (transitions from each consonant/vowel to the next). Adding 25% will double the speed of the whole voice synthesizer.

Trick!

You could be tempted to increase Phoneme Speed / Wave rate in higher tempo songs affecting the voice's intelligibility. Please refer to “Making the most out of chipspeech in a sequencer” (page 21).

The 'Words' Tab – Singing and Talking Voices



Text/X-Sampa (singing only) : Switch between editing lyrics in English (or Japanese hiragana / katakana) or in phonetic notation (using X-SAMPa - similar to IPA but doesn't require special characters).

Inflection level (talking only) : Increase or decrease how high and low talking mode inflections go. Setting this to 0% will give you a robotic effect (great for voices like Otto Mozer).

Move Up : Swaps current line and previous one.

Move Down : Swaps current line and next one.

Loop Line : Keep looping the current line instead of advancing through the text. You can automate this using MIDI CC79 or plug-in parameter automation.

Copy All : Copy whole lyrics to clipboard.

Paste All : Paste whole lyrics from clipboard.

Undo : Undo the last change to the lyrics.

Preset : Load preset lyrics to test your voice settings. You can also save your own presets.

Text : Lyrics sung by chipspeech. You can edit the text, or cue a line for the next note. If you click on the line number, the line will also be cued next time playback is restarted.

The 'Words' Tab – Circuit Bending Voices



Bank : Selects which word bank you play from.

Selected Entry : Selects which entry from the bank should be played on next note.

Use Original Data Pitch : Controls whether the pitch used comes from the played note, or from the original pitch data (SP0 and TMS only).

Loop Start : Offsets where the data starts playing from.

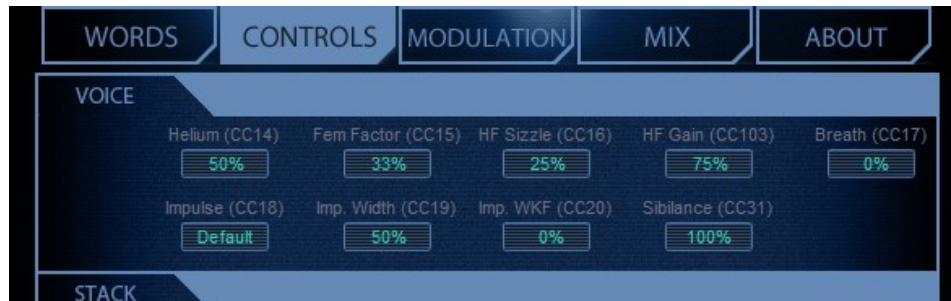
Loop Length : Lets you adjust how much data is played before the reading position is reset. If you set this to 0, you can scan through the data in real time by automating the Loop Start.

Loop : Selects whether notes loop or stop playing once the current entry was read.

Freeze Data : Loops the currently playing frame, emulating another famous circuit bending trick.

The 'Controls' Tab

Voice Parameters



Helium : This parameter scales the voice's resonance frequencies (formants). Raising this simulates a smaller mouth, giving rise to the "squirrel effect".

Fem Factor : This parameter modifies the voice's formants to be more male-like or female-like.

- Bert Gotrax, Dandy 704: 25% gives you the original male formants.
- Dee Klatt: 0%: husky male, 33%: higher male, 66%: female, 100%: higher female.

HF Sizzle : Controls the amount of top-end harmonics.

- Bert Gotrax, Dandy 704: stretches the high spectrum to increase coverage.
- Dee Klatt: controls the mix level of the built-in harmonic exciter.
- Other voices: Controls the gain of the high-end shelf filter.

HF Gain : Controls the high-end shelf filter on Bert Gotrax and Dandy 704 voices.

Breath : Controls the amount of noise added into the voice. (Bert Gotrax, Dandy 704 and Dee Klatt voices only)

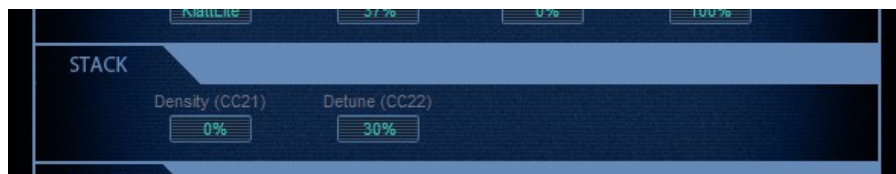
Impulse : Select which base waveform is used to simulate vocal cords. (no effect on Otto Mozer voice)

Impulse Width : Controls the stretch factor of the base waveform. On the square waveform, this controls pulse width. (no effect on Otto Mozer voice)

Impulse Width Key Follow : Controls how much the impulse width follows the keyboard. Setting this to 0% will use the same waveform for the whole keyboard. Setting this to 100% will fully scale the waveform as you play higher notes. (no effect on Otto Mozer voice)

Sibilance : Increase or decrease the volume of noisy and sibilant consonants (p, t, k, f, th, s, sh, h).

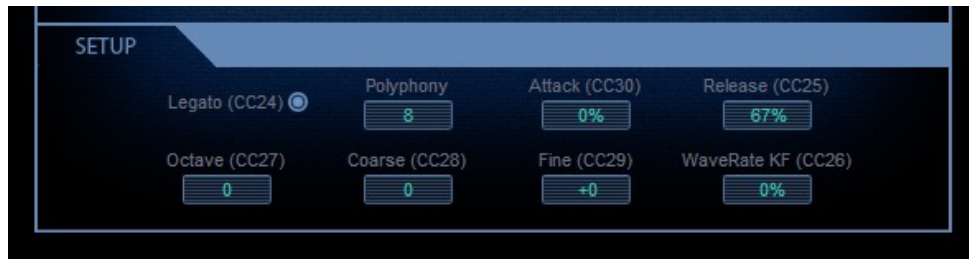
Stack Parameters



Density : Controls the number of detuned voices in the stack. 0% gives a single voice. 30~50% gives you about two voices. 100% gives you the full 9 voice stack.

Detune : Amount of detuning between the stack voices.

Setup Parameters



Legato : Activates legato mode. Lets you do legatos in the melody without triggering new syllables. (monophonic mode only)

Polyphony : Controls the maximum number of voices in polyphonic mode.

Attack : Controls how fast notes fade in when played. In monophonic mode, only applies to the first note of a group.

Release : Controls how fast the note will decay when released. The release is somewhat different in polyphonic mode.

Rate Key Follow : Controls how much the wave rate scales with the keyboard. 0% is fully pitch independent. 100% will fully scale speed, similar to a sampler.

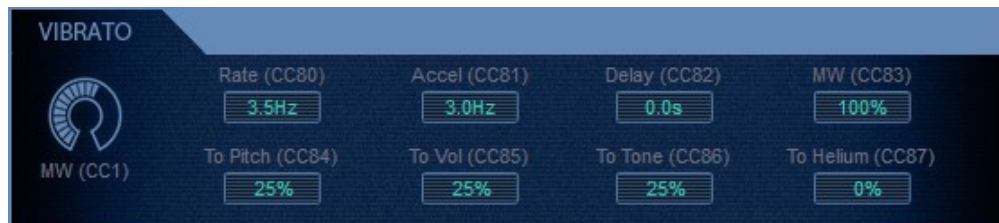
Octave : Increases/decreases the pitch in octave steps.

Coarse Tune : Increase/decrease the pitch in semitone steps.

Fine Tune : Increase/decrease the pitch in cents.

The 'Modulation' Tab

Vibrato Parameters



MW : Vibrato intensity, as controlled by the Mod Wheel .

Rate : How fast the vibrato is.

Acceleration : How much the vibrato rate increases when the vibrato amount is increased.

Delay : Delay time it takes for the vibrato to start.

Mod Wheel : How much effect the Mod Wheel (Midi CC1) has on vibrato amount.

To Pitch : How much vibrato is applied to pitch.

To Volume : How much vibrato is applied to volume (tremolo).

To Tone : How much vibrato is applied to impulse width (PWM LFO). (no effect on Otto Mozer voice)

To Helium : How much vibrato is applied to helium (formant LFO).

Expression Parameters



Exp : Expression level (MIDI CC11). This controls volume and lets you do crescendos/decrescendos.

Modwheel Amount : Expression level is controlled by CC11. This lets you control the expression level with the Mod Wheel (CC1) as well.

To Volume : How much effect expression has on volume.

To HF Sizzle : How much effect expression has on HF Sizzle / HF Gain.

Siz.>Fmts/Gain : Controls whether expression and velocity are applied on high frequency formant stretching (same as HF Sizzle parameter), or on high frequency gain (same as HF Gain parameter).

Vel to Vol : Controls the amount of effect velocity has on volume.

Vel to WaveRt : Controls the amount of effect velocity has on volume.

Vel to Glide : Controls the amount of effect velocity has on volume.

Vel to Sizzle : Controls the amount of effect velocity has on high frequency formant stretching or high frequency gain.

Humanize and Glide Parameters



Pitch : How much random variation is added to pitch.

Volume : How much random variation is added to volume.

Glide : How fast pitch changes when playing separate notes (triggering new syllables). (Monophonic mode only)

Legato Glide : How fast pitch changes when playing legato notes. In combination with legato mode, this lets you easily simulate voice inflections without using the pitch bender. (Monophonic+Legato mode only)

Humanity : Controls the shape of the glide, from highly robotic (0%) to more human-like (100%).

Volume Smooth : Controls whether volume/velocity tone changes happen sharply (0%) or smoothly (100%).

The 'Mix' Tab

FX Type : Selects between the Ambience effect and the Detune effect.

Effect Parameters – Ambience (reverb)



Ambience is a very smooth and flexible stereo reverb effect created by Magnus Jonsson, known mostly for its Ambience VST/AU version. In particular, the decay time can be shaped to different values in the low/mid/high frequency ranges.

Preset : Lets you load reverb settings from a preset.

Decay : How long the reverb decays.

Diffusion : How smooth the decay tail is.

Size : Time between the echos (simulates a small or large room).

Predelay : Amount of delay between the original sound and its reverb tail.

Width : Amount of stereo separation (how different the reverb is on the left vs right side).

Quality : Increases the complexity of the reverb tail, at the cost of requiring more CPU to process.

Output : Volume of reverb.

Equalizer Lo Freq : Frequency under which the low gain is applied.

Equalizer Lo Gain : Gain applied on low frequencies.

Equalizer Hi Freq : Frequency over which the high gain is applied.

Equalizer Hi Gain : Gain applied on high frequencies.

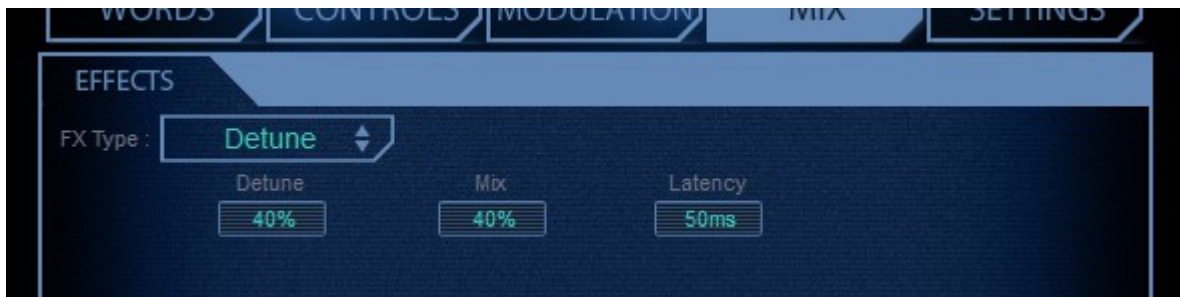
Damping Lo Freq : Frequency under which low damping is applied.

Damping Lo Amt : Decay time change under low damping frequency.

Damping Hi Freq : Frequency over which high damping is applied.

Damping Hi Amt : Decay time change over high damping frequency.

Effect Parameters – Detune



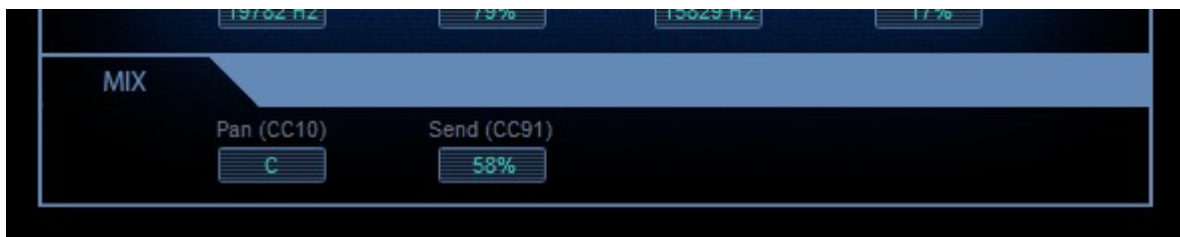
This effect creates two detuned versions of the voice input, one panned left and one panned right.

Detune : How detuned the simulated voices are.

Mix : How much effect is mixed in.

Latency : How delayed the simulated voices are.

Mix Parameters



Pan : Balance of the voice between the left and right channels.

Send : How much of the voice is sent to the effect. Setting this to 0% disables the effect.

Using chipspeech as a Standalone App

Plogue chipspeech can be launched by itself and played live via MIDI keyboard or other MIDI controller. The standalone version of Plogue chipspeech effectively makes your computer, audio hardware and MIDI keyboard into a virtual synthesizer that can be played independently of other programs. Unlike using it as a plug-in within a sequencer, your recording ability is limited and you can not edit your performance (though you can use various audio software programs for this).

Launching chipspeech in Standalone Mode

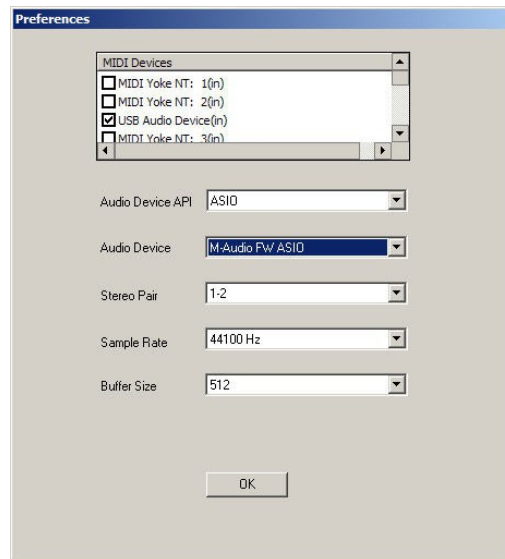
First, make sure that you have followed the instructions in the installation section of this manual. Be certain that your audio/sound interface and MIDI hardware interfaces are properly connected to the computer, your speakers or headphones are connected and everything is powered up.

To launch Plogue chipspeech as a standalone application, click on the Plogue chipspeech logo on your desktop or go to the Program Files or Applications folder and launch “chipspeech”.

Basic Setup Information for Standalone Mode

To use the standalone version you have to configure the Audio and MIDI settings in the Preferences dialog box (found in the Tools menu) before you can play. When used as a plug-in, the host sequencer or tracker program has already set up its audio and MIDI connections, and the Plogue chipspeech “plugs in” to them. However, with standalone operation Plogue chipspeech communicates directly with your audio and MIDI interface. Setup for Mac and Windows computers is similar, except where indicated. Note that if you change your audio interface, you will almost certainly need to readjust these settings.

Call up the Preferences setup dialog from the Tools menu on the Plogue chipspeech standalone interface. You'll see drop-down menus for MIDI Device, Audio Devices, Stereo Pair, Sample Rate and Buffer Size.



- ◆ **MIDI Device Menu:** All supported (and installed) MIDI interfaces are available in this drop-down list. Select the desired MIDI device from the list. The Plogue Aria Engine sends and receives MIDI on these selected devices.
- ◆ **Audio Device Menu:** All supported (and installed) audio interfaces are available in this drop-down list. Select the desired audio device from the list.
- ◆ **Stereo Pair:** Here you can define which of the stereo outputs should be used. Many pro audio devices have a variety of outputs, so you may choose which of these are available on your system that you would like chipspeech to output through.
- ◆ **Sample Rate:** Depending on the sound card and driver you are using, various sample rates are available. Set the

desired sample rate here.

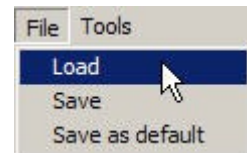
- ◆ **Buffer Size:** The buffer size setting will determine the delay between pressing a key on your MIDI keyboard and hearing the sound (a/k/a 'latency'). The default buffer size of 512 samples typically works well, but smaller buffer sizes will give a faster response (lower latency) and higher buffer sizes will give better audio performance (more polyphony and higher fidelity). Most modern computers and audio interfaces can handle a buffer size of 512 samples without a significant reduction in polyphony. If the sound is breaking up or crackling when a note sounds, then first check that the audio connections and wiring are good. Then, try a larger audio buffer size setting. Please note that there is typically a trade-off between higher buffer sizes (polyphony and sound fidelity) and lower buffer sizes (faster response or lower latency). Also note that the sound card buffer size settings determine latency, rather than Plogue chipspeech Player itself.

Once you have your Audio and MIDI set up, and have loaded one of the snapshots, you can begin playing chipspeech. Try playing a key on your MIDI keyboard. If the MIDI and audio configurations are correct, you should hear the corresponding synth note. If not, check the MIDI connections and wiring, and the MIDI output channel of your MIDI keyboard. Also check that the channel is specified correctly. If you are hearing the notes play, then the basic configuration is complete, and you are ready to use Plogue chipspeech.

File Menu for Loading and Saving Snapshots in the Standalone

Configuration presets (.ariax files) for chipspeech can be saved and loaded. This gives the user the ability to customize instruments setups to suit personal preferences and save configurations for convenient future use. The File menu choices are:

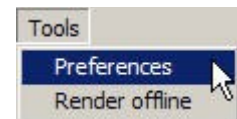
- ◆ **Load**—any saved configuration preset files in Aria format can be loaded by clicking on this choice and selecting the desired file.
- ◆ **Save**—any configuration can be saved by clicking on this choice, typing a name for the custom preset and saving to a desired location.
- ◆ **Save as default**—any settings can be saved as part of the default, to be loaded automatically at the time the chipspeech player is booted in standalone mode.



Tools Menu in Standalone ONLY

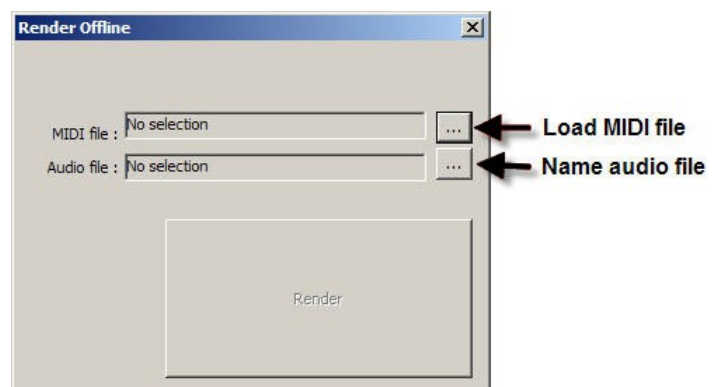
The Tools menu includes:

- ◆ **Preferences**—as described in the basic setup information above.
- ◆ **Render offline**—MIDI files can be rendered to audio offline using this feature.

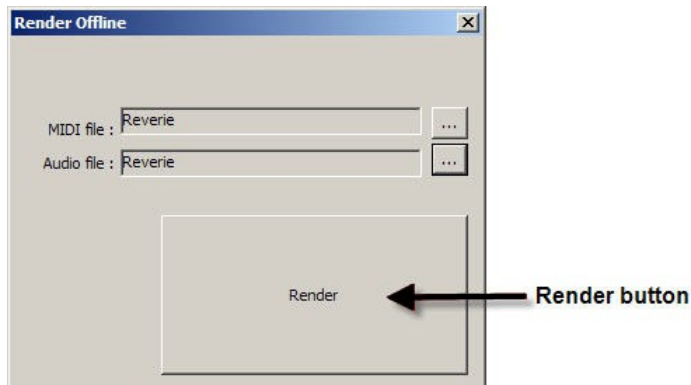


To use the Render Offline feature:

1. Click on the Load MIDI file button.
2. Select the desired MIDI file
3. Click on the Name audio file button.
4. Name the audio file and specify its location.



Once the files are in place, there is just one more step:



Click on the Render button. The audio file will be rendered to the specified location.

MIDI File Player and Audio Recorder (Standalone Version Only)



The standalone version of the player includes a MIDI file player and Audio Recorder. It is located at the bottom of the standalone window. This feature enables the user to load and play a standard MIDI file and render an audio file from it. The MIDI controls are located on the left hand side and the audio controls on the right.

To load and play a MIDI file:

1. Click on the “Load” button.
2. Choose the file you wish to load.
3. Click on the “Play” button.

The MIDI file will play back using the presently loaded chip setup and any other features (e.g. Ambience reverb) activated in the player. The progress bar will move to the right as the file is played. To start the file from the beginning, click on the “Go to beginning” button. The file can be stopped at any time by clicking on the “Stop” button.

To record your playback to an audio file:

1. Click on the “Name file” button.
2. Name the file and specify its desired location.
3. Click on the “Record” button.
4. Start playback by clicking on the MIDI “Play” button.
5. When playback finishes, click on the MIDI “Stop” button.
6. Click on the Audio “Stop” button.

The audio file will be located on your drive in the location you specified.

Note:

There is also an offline rendering feature (explained previously) that is located in the “Tools” menu of the standalone version.

Using chipspeech as a Plug-In

When used as a plug-in, Plogue chipspeech is not a standalone program, but rather a virtual instrument module that is seamlessly integrated into your favorite music software program or sequencer. They are called "plug-ins" because these are modular software applications that run inside a "host" music application, (e.g. a sequencer or tracker program, typically).

Using chipspeech as a plug-in has various advantages:

- ◆ MIDI recording and sequencing of Plogue chipspeech.
- ◆ Audio mixing of Plogue chipspeech with other instrument tracks within a single program.
- ◆ Easy automation of Plogue chipspeech parameters through the use of MIDI CC's or the host's automation.
- ◆ Effect processing of Plogue chipspeech using effect plug-ins in a music software program Saving and recalling of all plug-in settings when the music software program or sequencer file is reloaded.
- ◆ Integration with other instruments into a "virtual studio".

A great thing about plug-ins is that they work with a large variety of compatible music programs. For example, Plogue chipspeech can be used as a VST plug-in in many VST music programs, sequencers, supported tracker programs and hosts. It can also be used as an Audio Units, AAX or RTAS plug-in.

Plug-in Standard	Description	Windows 32bit	Windows 64bit	Mac 32bit	Mac 64bit
VST 	The VST plug-in stands for Virtual Studio Technology and was developed by Steinberg, the makers of the Cubase family of audio programs. It is also used by Cakewalk Sonar, Ableton Live, Renoise, Bidule, FL Studio and other sequencers.	X	X	X	X
Audio Units  Audio Units	The Audio Units (AU) plug-in standard was developed by Apple Inc. for Core Audio under Mac OS X. Audio Units is the preferred plug-in format on Mac OS X and is used by Apple GarageBand & Logic and MOTU Digital Performer.			X	X
AAX & RTAS 	AAX and RTAS plug-ins are designed to work in the Digidesign Pro Tools environment. Pro Tools hardware and software are used extensively in the pro audio and post production communities.	RTAS	AAX	RTAS	AAX

Plug-in Use

To use Plogue chipspeech as a plug-in instrument, you simply launch your host music application/sequencer first and then launch Plogue chipspeech from within it. Make sure that your sequencing host program is properly installed and configured, and that it is producing sound properly. Used as a plug-in, chipspeech's audio and MIDI data is managed by the host music software application.

Each music software application has its own approach to handling plug-in instruments. They each have a different method of installation as well as differing means of loading and accessing plug-ins. It is important to make sure that you refer to the instructions in your music software application's manual regarding the loading and operation of plug-in instruments.

Windows VST Setup

Make sure the chipspeech plugin is within the VST plugin folder used by your host, so that it can be scanned and recognized on startup.

Making the most out of chipspeech in a sequencer

If you are used to the way other “Standalone” vocal synthesizers work, you will probably enter MIDI data like the following for détaché'd (non-legato) syllables:

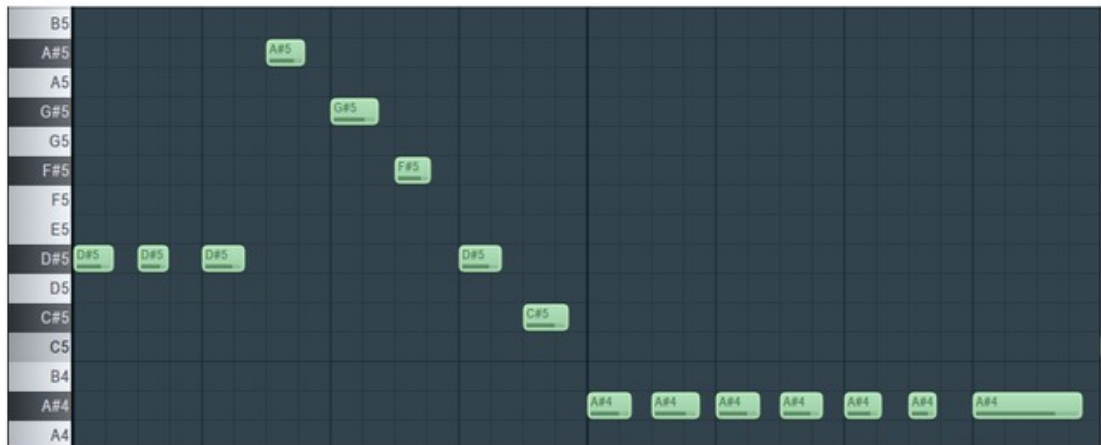


Doing so will sadly not generate the results you expect here.

Since chipspeech is a real time plug-in and not a “timeline controlling” application, it is not aware of the content of the “whole” song in advance (it *cannot* read into the future). And because generated transitions between syllables start on note releases, if a new note starts *just after* the end of the previous one, a very quick transition will be inserted by chipspeech before the new note starts, and so the whole vocal line will drift in the future.

Even though you can speed up transitions by setting the **Phoneme Speed** and **Wave Rate** parameters to extreme values, this will only make the overall voice sound more robotic.

Instead you should make each note a little shorter than you would normally expect them to last, between 50-75% of its intended length. The results should look something like this:



If you listen carefully you will realize that different note lengths work better for different words, or songs styles, and you may start to use this 'restriction' creatively. For instance, you can emulate spoken word if you use very short notes.

However, if you are entering notes using live MIDI input from keyboard, then you don't have to worry as you will unconsciously figure this out just by "feel" or intuition.

A complete guide to chipspeech

In order to make **legato** sections (changing a note without advancing the syllable) however requires overlaps.

As an example, typing in “I Am” in the text and playing the following:



should give you “I-i-i...A-a-a-am”.

You can of course mix and match both methods in a song, as a normal singer would.

You may use a common technique that consists in recording your own voice singing what you have in mind, and temporarily place it side by side with the virtual singer, in order to give you hints of how to encode a particular section.

Use your ears and experiment! Good luck!

Plogue chipspeech[®]

Copyright © 2015-2018 Plogue Art et Technologie, Inc. All rights reserved.

Produced by :	David Viens
ARIA Engine Design/Development :	Plogue Art et Technologie, Inc.
Voice Models :	David Viens Hubert Lamontagne
Sound Programming :	David Viens Eric Patenaude Hubert Lamontagne
Hardware research and modeling :	David Viens
Phoneme Sequencer and Mangling :	Hubert Lamontagne
In-house Testing and QA :	Eric Patenaude
Ambience Reverb :	Magnus Jonsson
User interface :	Matt Galais Eric Patenaude
Art :	Matt Galais

Voice Credits

Bert Gotrax	Was built from analysis of sound recordings of a Votrax SC-01. Recordings copyrighted Plogue Art et Technologie, Inc.
CiderTalk'84	Derived from: SoftVoice Text-to-Speech Engine Copyright 1984-2015, SoftVoice, Inc. All Rights Reserved
Dandy 704	Was built from analysis of sound recordings courtesy of Alcatel-Lucent / Bell Labs (C) [1963] Alcatel-Lucent. All Rights Reserved.
Dee Klatt	Was built from Dennis H. Klatt's synthesizer originally published in Journal Acoustic Society of America, Mar. 1980.
Lady Parsec(HD and LPC10)	Was built from analysis of sound recordings copyrighted Plogue Art et Technologie, Inc.
Otto Mozer	Was built from analysis of data files copyrighted Dr. Forrest S. Mozer. Used with kind permission.
Rotten.ST	Original voice by Martin Day and Andy Beveridge. Additional code by Stefan Stenzel.
Spencer AL2	Was built from analysis of ROM data and algorithms copyrighted Microchip, Inc.
Terminal 99	Was built with ROM data and algorithms copyrighted Texas Instruments Incorporated. Distributed under license from Texas Instruments (TI).
TMS Bending	Alpiner, Moon Mine, Parsec, Terminal Emulator II and Speech Module ROM data is used under license from Texas Instruments (TI). TI makes no warranty with respect to the programs and is under no obligation to provide any support or assistance with respect to the programs. No liability is accepted on the part of Texas Instruments or the author with respect to use, copying or distribution of the programs.

Acknowledgments

Producing Plogue chipspeech would not have been possible without the combined help, talent and support of many extraordinary people. We are grateful to those who have contributed and would like to thank them all.

Matt Gallais
Dr. Forrest S. Mozer
Gene A Frantz
Edward J Eckert
Herbert W. Foster
Chibi-Tech
Jonathan Gevanyahu
Kevin Horton
Aubree Anderson
Martine Drapeau
aliceffekt
Crusher-P

THANK YOU!

Appendix : X-Sampa reference

Plogue chipspeech uses X-SAMPA for phonetic transcription. This was chosen to eventually allow for multiple language support (unlike Arpabet or Votox), and to remove the need for non-ASCII character support (which is a problem with IPA and Americanist notation).

Quick chart

Special

Pause	#
Short	\$1..\$4
Long	\$6..\$9
Variant	*1..*9

Consonants

	Labial		Dental		Alveolar		Palatal		Velar		Glottal
Stop	p	b			t	d	tS	dZ	k	g	
Fricative	f	v	T	D	s	z	S	Z			h
Nasal	m				n				N		
Liquid	w				l	r\	j				

Vowels

	Front	Central		Back
High	i			u
Near-high	ɪ			ʊ
Mid-high	eɪ			oʊ
Mid-low	E	V	@	O
Low	{			A
Diphthong	AI	oɪ		{U
High -r	ɪr\			ʊr\
Mid -r	E r\	@`		o r\
Low -r				A r\

Full list of phonemes

X-Sampa	IPA	Spellings	Example
Special			
#		.	[Pause, no sound]
[space]	.		[Syllable break]
\$6 ... \$9	:		[Longer]
\$1 ... \$4			[Shorter]
*1 ... *9			[Variant phoneme (if available)]
Consonants			
p	p	p	Pin
t	t	t	Tin
tS	tʃ	ch	CHin
k	k	k, q, c (hard)	Kin
b	b	b	Bin
d	d	d	Din
dZ	dʒ	j, ge, gi (soft g)	Gin
g	g	g (hard g), gu	Grin
f	f	f, ph	Fin
T	θ	th (hard)	THin
s	s	s (hard), ss, sc, ce, ci	Sin
S	ʃ	sh	SHin
h	h	h	Him
v	v	v	Video
D	ð	th (soft)	THis
z	z	z, s (soft)	Zit
Z	ʒ	zh, ge (french), su, zu	garaGe, pleaSure, aZure, ZHukhov
m	m	m	kiM
n	n	n	kiN
N	ŋ	ng	kiNG
w	w	w	Wing
l	l	l	Limb
r\	ɹ	r	Rim
j	j	y	Yard

Vowels			
i	i	ee, ea	bEAn
l	ɪ	i (short)	bln
el	eɪ	a (long)	bAne
E	ɛ	e	bEn
{	æ	a (short)	bAn
V	ʌ	u (short)	bUn
@	ə	Unstressed vowels (e, u...)	thE
u	u	oo	sOOn
U	ʊ	ou (short), oo	shOUld
oU	oʊ	oa, o	bOne
O	ɔ	a, o	bAll
A	ɑ	a, o	PAIm
Al	ɑɪ	i (long)	blnd
ol	oɪ	oy	bOY
{U	æʊ	ou (long)	bOUnd
Ir\	ɪ	eer, ear	bEER
Er\	ɛ	ere, ear	bEAR
@`	ə	er, ur, ir	bURn
Ur\	ʊ	oor	pOOR
or\	oʊ	or, oar	bOAR
Ar\	ɑ	ar	bAR
Japanese-specific			
e	e	e	
a	a	a	
o	o	o	
M	ɯ	u	
4	r	r	
Miscellaneous			
R	ʁ	r (French, German)	
y	y	French u, German ü	
Y	ʏ	German ü (short)	
2	ø	French eu, German ö	

A complete guide to chipspeech

6	ɐ	German er	
B	β	Spanish soft b	
G	ɣ	Spanish soft g	
C	ç	German ch (after ieäöü)	
x	x	Spanish j (Americas)	
X	χ	German ch (after aou)	
l*1	l	Non-velar L (French, German, Spanish...)	
L	ʎ	Spanish LL	
H	ɥ	French u in ui	
J	ɲ	Ñ, French gn	

Alphabetic tables

Consonants

b	b	g (soft)	dZ	q	k	th	T, D
c (hard)	k	h	h	qu	kw	-tial	S@l
c (soft)	s	j	dZ	r	r\	-tion, -tious	S@n, S@s
-cial	S@l	k	k	s	s, z	-ture	tS@`
-cion, -cious	S@n, S@s	l	l	-sion, -sian	Z@n	v	v
ch	tS	m	m	ss	s	w, wh	w
d	d	n	n	-ssion, -ssian	S@n	x	ks, gz
-dure	dZ@`	ng	N	sh	S	y	j
f	f	p	p	-sure	S@`, Z@`	z	Z
g (hard)	g	ph	f	t	t		

Vowels

a (long)	El	e (short)	E	ie	Al, i	ow	{U, oU
a (short)	{, A, O	e (unstressed)	@	o (long)	oU	u (long)	ju, u
a (unstressed)	@	ea, ee, ei	i	o (short)	O, A	u (short)	V, u
ah	A	ew	u, ju	o (unstressed)	@	u (unstressed)	@
ai	el	ey	el	oa	oU	ue	ju
au, aw	O	i (long)	Al, i	oo	u	uy	Al
ay	el	i (short)	l	ou	{U, U, oU	y	Al, l, i
e (long)	i, el	i (unstressed)	l, @	Oi, oy	ol		

Vowels +r

air	Er\	eer	Ir\	-ire	A j@`	ur	@`
ar	Ar\	eir	Er\	or, oar	or\	-ure	Ur\, jUr\
-are	Er\	er	@`	oor	Ur\	yr	Ir\, @`
aur	or\	-ere	Er\, Ir\	our	or\, Ur\, { w@`		
ear	Ir\, Er\	ir	@`	oir	wA j@`		